

RECEIVED

JUL 19 2001

TECH CENTER 1600/2900

1644

PH#2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/019,348A

DATE: 06/29/2001

TIME: 11:35:43

Input Set: A:\10287-031001.TXT

Output Set: N:\CRF3\06292001\I019348A.raw

P.S

ENTERED

4 <110> APPLICANT: Georgopoulos, Katia
 5 Morgan, Bruce A.
 7 <120> TITLE OF INVENTION: AIOLOS GENE
 10 <130> FILE REFERENCE: 10287-031001
 12 <140> CURRENT APPLICATION NUMBER: US 09/019,348A
 13 <141> CURRENT FILING DATE: 1998-02-05
 15 <150> PRIOR APPLICATION NUMBER: US 08/733,622
 16 <151> PRIOR FILING DATE: 1996-10-17
 18 <150> PRIOR APPLICATION NUMBER: US 60/017,646
 19 <151> PRIOR FILING DATE: 1996-05-14
 21 <150> PRIOR APPLICATION NUMBER: US 60/005,529
 22 <151> PRIOR FILING DATE: 1995-10-18
 24 <160> NUMBER OF SEQ ID NOS: 38
 26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 28 <210> SEQ ID NO: 1
 29 <211> LENGTH: 1984
 30 <212> TYPE: DNA
 31 <213> ORGANISM: Mus musculus
 33 <220> FEATURE:
 34 <221> NAME/KEY: CDS
 35 <222> LOCATION: (374)...(1894)
 37 <400> SEQUENCE: 1
 38 cacgagcgca caccgctcgg ctctccttgc gacacgccct catccccggt gtttctcaag 60
 39 tagacgtccc gagacggtcg ctgaggcact gtttccacgc gatcagggtt cctcagggtt 120
 40 gacattcaaa agtgggtgcg gaacccgcg cactcggagc gtgctttaa gcgcccgcca 180
 41 gccagcgccg ctctaaccct gcgccccggc tgccggcgcc tccccgctg catctgcgcc 240
 42 gacgcgaccg agcgatccc gggcctccct gcgcccggaa tctcccgcca gccgcgcggg 300
 43 tccccacggc agcagcacgt ggagcggccg cggagcctga gcgacagctg cagcccgccg 360
 44 ggcccgcgcc gac atg gaa gat ata caa ccg act gtg gag ctg aaa agc 409
 45 Met Glu Asp Ile Gln Pro Thr Val Glu Leu Lys Ser
 46 1 5 10
 48 acg gag gag cag cct ctg ccc aca gag agc cca gac gct ctg aat gac 457
 49 Thr Glu Glu Gln Pro Leu Pro Thr Glu Ser Pro Asp Ala Leu Asn Asp
 50 15 20 25
 52 tac agc ttg ccc aaa cct cat gag ata gaa aac gtg gac agt aga gaa 505
 53 Tyr Ser Leu Pro Lys Pro His Glu Ile Glu Asn Val Asp Ser Arg Glu
 54 30 35 40
 56 gcc cca gcc aat gaa gac gaa gat gca gga gaa gat tcg atg aaa gtg 553
 57 Ala Pro Ala Asn Glu Asp Glu Asp Ala Gly Glu Asp Ser Met Lys Val
 58 45 50 55 60
 60 aaa gat gaa tac agc gac aga gat gag aac att atg aag ccg gag ccc 601
 61 Lys Asp Glu Tyr Ser Asp Arg Asp Glu Asn Ile Met Lys Pro Glu Pro
 62 65 70 75
 64 atg gga gat gca gaa gag agt gaa atg cct tac agc tat gca aga gaa 649
 65 Met Gly Asp Ala Glu Glu Ser Glu Met Pro Tyr Ser Tyr Ala Arg Glu
 66 80 85 90
 68 tac agc gac tat gaa agc att aag ctg gag aga cac gtg ccc tat gac 697

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/019,348A

TIME: 11:35:43

Input Set : A:\10287-031001.TXT

Output Set: N:\CRF3\06292001\I019348A.raw

69	Tyr	Ser	Asp	Tyr	Glu	Ser	Ile	Lys	Leu	Glu	Arg	His	Val	Pro	Tyr	Asp	
70			95					100					105				
72	aac	agc	aga	cca	acc	agt	ggg	aag	atg	aac	tgc	gac	gtg	tgc	ggg	tta	745
73	Asn	Ser	Arg	Pro	Thr	Ser	Gly	Lys	Met	Asn	Cys	Asp	Val	Cys	Gly	Leu	
74		110					115					120					
76	tcc	tgc	att	agc	ttc	aac	gtc	ttg	atg	gtt	cat	aag	cga	agc	cat	acc	793
77	Ser	Cys	Ile	Ser	Phe	Asn	Val	Leu	Met	Val	His	Lys	Arg	Ser	His	Thr	
78	125					130					135					140	
80	ggc	gaa	cgc	ccg	ttc	cag	tgt	aat	cag	tgc	ggg	gca	tct	ttt	act	cag	841
81	Gly	Glu	Arg	Pro	Phe	Gln	Cys	Asn	Gln	Cys	Gly	Ala	Ser	Phe	Thr	Gln	
82					145					150					155		
84	aaa	ggt	aac	ctc	ctc	cgt	cat	att	aaa	ctg	cac	acg	ggg	gaa	aaa	cct	889
85	Lys	Gly	Asn	Leu	Leu	Arg	His	Ile	Lys	Leu	His	Thr	Gly	Glu	Lys	Pro	
86				160					165					170			
88	ttt	aag	tgt	cac	ctc	tgc	aac	tac	gca	tgc	caa	agg	aga	gat	gcg	ctc	937
89	Phe	Lys	Cys	His	Leu	Cys	Asn	Tyr	Ala	Cys	Gln	Arg	Arg	Asp	Ala	Leu	
90		175					180					185					
92	acg	gga	cac	ctt	agg	aca	cat	tct	gtg	gag	aag	ccg	tac	aag	tgt	gag	985
93	Thr	Gly	His	Leu	Arg	Thr	His	Ser	Val	Glu	Lys	Pro	Tyr	Lys	Cys	Glu	
94		190				195					200						
96	ttc	tgc	gga	aga	agc	tac	aag	cag	aga	agc	tcc	ctg	gag	gag	cac	aag	1033
97	Phe	Cys	Gly	Arg	Ser	Tyr	Lys	Gln	Arg	Ser	Ser	Leu	Glu	Glu	His	Lys	
98	205				210				215					220			
100	gaa	cgc	tgc	cga	gct	ttt	ctt	cag	aac	cct	gac	ctg	ggg	gac	gct	gca	1081
101	Glu	Arg	Cys	Arg	Ala	Phe	Leu	Gln	Asn	Pro	Asp	Leu	Gly	Asp	Ala	Ala	
102				225					230				235				
104	agt	gtg	gag	gca	aga	cac	atc	aaa	gcc	gag	atg	gga	agt	gag	aga	gct	1129
105	Ser	Val	Glu	Ala	Arg	His	Ile	Lys	Ala	Glu	Met	Gly	Ser	Glu	Arg	Ala	
106				240				245				250					
108	ctc	gtc	ctg	gac	aga	tta	gca	agc	aat	gtg	gct	aag	cga	aaa	agc	tcg	1177
109	Leu	Val	Leu	Asp	Arg	Leu	Ala	Ser	Asn	Val	Ala	Lys	Arg	Lys	Ser	Ser	
110			255				260					265					
112	atg	cct	cag	aaa	ttc	atc	ggt	gag	aag	cgg	cac	tgc	ttc	gat	gcc	aac	1225
113	Met	Pro	Gln	Lys	Phe	Ile	Gly	Glu	Lys	Arg	His	Cys	Phe	Asp	Ala	Asn	
114		270				275					280						
116	tac	aat	ccc	ggc	tac	atg	tac	gag	aag	gag	aac	gag	atg	atg	cag	acc	1273
117	Tyr	Asn	Pro	Gly	Tyr	Met	Tyr	Glu	Lys	Glu	Asn	Glu	Met	Met	Gln	Thr	
118	285					290					295					300	
120	cgg	atg	atg	gac	caa	gcc	atc	aat	aac	gcc	atc	agc	tat	cta	ggg	gct	1321
121	Arg	Met	Met	Asp	Gln	Ala	Ile	Asn	Asn	Ala	Ile	Ser	Tyr	Leu	Gly	Ala	
122				305					310					315			
124	gaa	gcc	ttc	cgc	ccc	tta	gtc	cag	act	ccg	cct	gct	ccc	acc	tct	gag	1369
125	Glu	Ala	Phe	Arg	Pro	Leu	Val	Gln	Thr	Pro	Pro	Ala	Pro	Thr	Ser	Glu	
126				320				325				330					
128	atg	gtc	cca	gtc	atc	agc	agt	gtg	tac	ccc	ata	gca	ctt	act	cgg	gcc	1417
129	Met	Val	Pro	Val	Ile	Ser	Ser	Val	Tyr	Pro	Ile	Ala	Leu	Thr	Arg	Ala	
130			335				340					345					
132	gat	atg	cca	atg	ggg	gcc	ccg	cag	gag	atg	gaa	aag	aaa	cgg	atc	ctc	1465
133	Asp	Met	Pro	Met	Gly	Ala	Pro	Gln	Glu	Met	Glu	Lys	Lys	Arg	Ile	Leu	

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/019,348A

TIME: 11:35:43

Input Set : A:\10287-031001.TXT

Output Set: N:\CRF3\06292001\I019348A.raw

```

134      350      355      360
136 ctg cca gag aag atc ttg cct tct gaa cga ggt ctg tcc ccc aat aac      1513
137 Leu Pro Glu Lys Ile Leu Pro Ser Glu Arg Gly Leu Ser Pro Asn Asn
138 365      370      375      380
140 agt gcc cag gac tcc aca gac acc gac agc aac cac gag gat cgc caa      1561
141 Ser Ala Gln Asp Ser Thr Asp Thr Asp Ser Asn His Glu Asp Arg Gln
142      385      390      395
144 cat ctc tac cag caa agc cac gtg gtc ctc ccc cag gcc cgc aat ggg      1609
145 His Leu Tyr Gln Gln Ser His Val Val Leu Pro Gln Ala Arg Asn Gly
146      400      405      410
148 atg cct ctt ctg aag gag gtc cct cgc tct ttt gaa ctc ctc aag ccc      1657
149 Met Pro Leu Leu Lys Glu Val Pro Arg Ser Phe Glu Leu Leu Lys Pro
150      415      420      425
152 cct ccc atc tgc ctg agg gac tcc atc aaa gtg atc aac aaa gaa ggg      1705
153 Pro Pro Ile Cys Leu Arg Asp Ser Ile Lys Val Ile Asn Lys Glu Gly
154      430      435      440
156 gag gtg atg gat gtg ttt cga tgt gac cac tgc cac gtc ctc ttc cta      1753
157 Glu Val Met Asp Val Phe Arg Cys Asp His Cys His Val Leu Phe Leu
158 445      450      455      460
160 gat tat gtg atg ttc acc atc cac atg ggg tgc cat ggt ttc cgt gat      1801
161 Asp Tyr Val Met Phe Thr Ile His Met Gly Cys His Gly Phe Arg Asp
162      465      470      475
164 ccc ttt gag tgt aac atg tgt ggc tat cga agc cac gat cgc tat gag      1849
165 Pro Phe Glu Cys Asn Met Cys Gly Tyr Arg Ser His Asp Arg Tyr Glu
166      480      485      490
168 ttc tcc tct cac atc gcc aga gga gag cac aga gcc atg ttg aag      1894
169 Phe Ser Ser His Ile Ala Arg Gly Glu His Arg Ala Met Leu Lys
170      495      500      505
172 tgagcatctg tcctcaatgc gaggggtcaac attgtttttt aaagctgatg gtagccttat      1954
173 ccagtagact gaactcaaac ccacctcgag      1984
175 <210> SEQ ID NO: 2
176 <211> LENGTH: 507
177 <212> TYPE: PRT
178 <213> ORGANISM: Mus musculus
180 <400> SEQUENCE: 2
181 Met Glu Asp Ile Gln Pro Thr Val Glu Leu Lys Ser Thr Glu Glu Gln
182 1      5      10      15
183 Pro Leu Pro Thr Glu Ser Pro Asp Ala Leu Asn Asp Tyr Ser Leu Pro
184      20      25      30
185 Lys Pro His Glu Ile Glu Asn Val Asp Ser Arg Glu Ala Pro Ala Asn
186      35      40      45
187 Glu Asp Glu Asp Ala Gly Glu Asp Ser Met Lys Val Lys Asp Glu Tyr
188      50      55      60
189 Ser Asp Arg Asp Glu Asn Ile Met Lys Pro Glu Pro Met Gly Asp Ala
190 65      70      75      80
191 Glu Glu Ser Glu Met Pro Tyr Ser Tyr Ala Arg Glu Tyr Ser Asp Tyr
192      85      90      95
193 Glu Ser Ile Lys Leu Glu Arg His Val Pro Tyr Asp Asn Ser Arg Pro
194      100      105      110

```

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/019,348A

TIME: 11:35:43

Input Set : A:\10287-031001.TXT

Output Set: N:\CRF3\06292001\I019348A.raw

```

195 Thr Ser Gly Lys Met Asn Cys Asp Val Cys Gly Leu Ser Cys Ile Ser
196      115      120      125
197 Phe Asn Val Leu Met Val His Lys Arg Ser His Thr Gly Glu Arg Pro
198      130      135      140
199 Phe Gln Cys Asn Gln Cys Gly Ala Ser Phe Thr Gln Lys Gly Asn Leu
200 145      150      155      160
201 Leu Arg His Ile Lys Leu His Thr Gly Glu Lys Pro Phe Lys Cys His
202      165      170      175
203 Leu Cys Asn Tyr Ala Cys Gln Arg Arg Asp Ala Leu Thr Gly His Leu
204      180      185      190
205 Arg Thr His Ser Val Glu Lys Pro Tyr Lys Cys Glu Phe Cys Gly Arg
206      195      200      205
207 Ser Tyr Lys Gln Arg Ser Ser Leu Glu Glu His Lys Glu Arg Cys Arg
208      210      215      220
209 Ala Phe Leu Gln Asn Pro Asp Leu Gly Asp Ala Ala Ser Val Glu Ala
210 225      230      235      240
211 Arg His Ile Lys Ala Glu Met Gly Ser Glu Arg Ala Leu Val Leu Asp
212      245      250      255
213 Arg Leu Ala Ser Asn Val Ala Lys Arg Lys Ser Ser Met Pro Gln Lys
214      260      265      270
215 Phe Ile Gly Glu Lys Arg His Cys Phe Asp Ala Asn Tyr Asn Pro Gly
216      275      280      285
217 Tyr Met Tyr Glu Lys Glu Asn Glu Met Met Gln Thr Arg Met Met Asp
218      290      295      300
219 Gln Ala Ile Asn Asn Ala Ile Ser Tyr Leu Gly Ala Glu Ala Phe Arg
220 305      310      315      320
221 Pro Leu Val Gln Thr Pro Pro Ala Pro Thr Ser Glu Met Val Pro Val
222      325      330      335
223 Ile Ser Ser Val Tyr Pro Ile Ala Leu Thr Arg Ala Asp Met Pro Met
224      340      345      350
225 Gly Ala Pro Gln Glu Met Glu Lys Lys Arg Ile Leu Leu Pro Glu Lys
226      355      360      365
227 Ile Leu Pro Ser Glu Arg Gly Leu Ser Pro Asn Asn Ser Ala Gln Asp
228      370      375      380
229 Ser Thr Asp Thr Asp Ser Asn His Glu Asp Arg Gln His Leu Tyr Gln
230 385      390      395      400
231 Gln Ser His Val Val Leu Pro Gln Ala Arg Asn Gly Met Pro Leu Leu
232      405      410      415
233 Lys Glu Val Pro Arg Ser Phe Glu Leu Lys Pro Pro Pro Ile Cys
234      420      425      430
235 Leu Arg Asp Ser Ile Lys Val Ile Asn Lys Glu Gly Glu Val Met Asp
236      435      440      445
237 Val Phe Arg Cys Asp His Cys His Val Leu Phe Leu Asp Tyr Val Met
238      450      455      460
239 Phe Thr Ile His Met Gly Cys His Gly Phe Arg Asp Pro Phe Glu Cys
240 465      470      475      480
241 Asn Met Cys Gly Tyr Arg Ser His Asp Arg Tyr Glu Phe Ser Ser His
242      485      490      495
243 Ile Ala Arg Gly Glu His Arg Ala Met Leu Lys

```

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/019,348A

TIME: 11:35:43

Input Set : A:\10287-031001.TXT

Output Set: N:\CRF3\06292001\I019348A.raw

```

244          500          505
246 <210> SEQ ID NO: 3
247 <211> LENGTH: 23
248 <212> TYPE: DNA
249 <213> ORGANISM: Artificial Sequence
251 <220> FEATURE:
252 <223> OTHER INFORMATION: primer for PCR
254 <400> SEQUENCE: 3
255 tayaccatyc acatgggctr cca 23
257 <210> SEQ ID NO: 4
258 <211> LENGTH: 21
259 <212> TYPE: DNA
260 <213> ORGANISM: Artificial Sequence
262 <220> FEATURE:
263 <223> OTHER INFORMATION: primer for PCR
265 <400> SEQUENCE: 4
266 rccrcacatg ttrcactyra a 21
268 <210> SEQ ID NO: 5
269 <211> LENGTH: 24
270 <212> TYPE: DNA
271 <213> ORGANISM: Artificial Sequence
273 <220> FEATURE:
274 <223> OTHER INFORMATION: primer for PCR
276 <400> SEQUENCE: 5
277 gtgtgcgggt tatectgcat tagc 24
279 <210> SEQ ID NO: 6
280 <211> LENGTH: 24
281 <212> TYPE: DNA
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: primer for PCR
287 <400> SEQUENCE: 6
288 atcgaagcag tgccgcttct cacc 24
290 <210> SEQ ID NO: 7
291 <211> LENGTH: 628
292 <212> TYPE: DNA
293 <213> ORGANISM: Homo sapiens
295 <220> FEATURE:
296 <221> NAME/KEY: CDS
297 <222> LOCATION: (1)...(627)
299 <400> SEQUENCE: 7
300 gaa aga gat gag aat gtt tta aag tca gaa ccc atg gga aat gca gaa 48
301 Glu Arg Asp Glu Asn Val Leu Lys Ser Glu Pro Met Gly Asn Ala Glu
302 1 5 10 15
304 gag cct gaa atc cct tac agc tat tca aga gaa tat aat gaa tat gaa 96
305 Glu Pro Glu Ile Pro Tyr Ser Tyr Ser Arg Glu Tyr Asn Glu Tyr Glu
306 20 25 30
308 aac att aag ttg gag aga cat gtt gtc tca ttc gat agt agc agg cca 144
309 Asn Ile Lys Leu Glu Arg His Val Val Ser Phe Asp Ser Ser Arg Pro

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

file://C:\Cr3\Outhold\Vsri019348A.htm

6/29/01

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/019,348A

DATE: 06/29/2001

TIME: 11:35:44

Input Set : A:\10287-031001.TXT

Output Set: N:\CRF3\06292001\I019348A.raw

L:1316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1324 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1336 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1338 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1340 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1344 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1346 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1348 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1350 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1352 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1354 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1356 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1358 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1364 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1366 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1368 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1374 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:1984 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
L:1986 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
L:1988 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
L:2002 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
L:2004 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37